

# Phylogenetic Analysis of a Novel Molecular Isolate of Spotted Fever Group Rickettsiae from Northern Peru

## *Candidatus* Rickettsia andeanae

JU JIANG,<sup>a,b</sup> PATRICK J. BLAIR,<sup>c</sup> VIDAL FELICES,<sup>c</sup> CECILIA MORON,<sup>d</sup> MANUEL CESPEDES,<sup>d</sup> ELIZABETH ANAYA,<sup>d</sup> GEORGE B. SCHOELEER,<sup>c</sup> JOHN W. SUMNER,<sup>e</sup> JAMES G. OLSON,<sup>c</sup> AND ALLEN L. RICHARDS<sup>a,f</sup>

<sup>a</sup>Naval Medical Research Center Silver Spring, Maryland, 20910, USA

<sup>b</sup>Henry M. Jackson Foundation, Rockville, Maryland, 20852, USA

<sup>c</sup>Naval Medical Research Center Detachment Lima, Peru

<sup>d</sup>National Institute of Health, Ministry of Health, Lima, Peru

<sup>e</sup>Centers for Disease Control and Prevention, Atlanta, Georgia, 30033, USA

<sup>f</sup>Uniformed Services University, Bethesda, Maryland, 20814, USA

**ABSTRACT:** Phylogenetic analysis of five rickettsial genes (17-kDa gene, *gltA*, *ompB*, *ompA*, and *sca4*) from two molecular isolates of *Candidatus* Rickettsia andeanae from two ticks (*Amblyomma maculatum* and *Ixodes boliviensis*) collected from two domestic horses living in two separate locations in northern Peru (Coletas and Naranjo) was conducted to more clearly characterize this recently reported novel spotted fever group (SFG) rickettsia. Following nested polymerase chain reaction (PCR) amplification of the 17-kDa gene, *gltA*, *ompB*, *ompA*, and *sca4*, amplicons were purified, sequenced, and compared to those downloaded from GenBank. Phylogenetic analyses of the *Candidatus* Rickettsia andeanae sequences generated from 17-kDa gene (483 bp), *gltA* (1185 bp), *ompA* (1598 bp), *ompB* (4839 bp), and *sca4* (2634 bp) demonstrated that they aligned strongly with those of SFG rickettsiae. Moreover, the sequences of these five genes most closely aligned with the following rickettsiae: *ompA*: *Rickettsia* sp RpA4 (98.03%), *R. sp* DnS28 (97.90%), and *R. rhipicephali* and *R. massiliae* (97.11%); *ompB*: *R. aeschlimannii* (97.22%), *R. rhipicephali* (97.20%), and *R. sp* Bar 29 (97.10%); and *sca4*: *R. massiliae* (97.8%), *R. rhipicephali*, and *R. slovacae* (97.7%). These results from the additional phylogenetic analyses of *Candidatus* Rickettsia andeanae confirm its inclusion within, and distance and uniqueness from, other known SFG rickettsiae.

**KEYWORDS:** Peru; ticks; spotted fever rickettsiae; phylogenetic analysis

Address for correspondence: Allen L. Richards, Ph.D., Director, Rickettsial Diseases Department, Naval Medical Research Center, 503 Robert Grant Avenue, Silver Spring, MD 20910-7500. Voice: (301) 319-7668; fax: (301) 319-9458.  
RichardsA@nmrc.navy.mil

Ann. N.Y. Acad. Sci. 1063: 337–342 (2005). © 2005 New York Academy of Sciences.  
doi: 10.1196/annals.1355.054

Report Documentation Page				Form Approved OMB No. 0704-0188	
Public reporting burden for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.					
1. REPORT DATE <b>2005</b>		2. REPORT TYPE <b>N/A</b>		3. DATES COVERED <b>-</b>	
4. TITLE AND SUBTITLE <b>Phylogenetic Analysis of a Novel Molecular Isolate of Spotted Fever Group Rickettsiae from Northern Peru</b>				5a. CONTRACT NUMBER	
				5b. GRANT NUMBER	
				5c. PROGRAM ELEMENT NUMBER	
6. AUTHOR(S)				5d. PROJECT NUMBER	
				5e. TASK NUMBER	
				5f. WORK UNIT NUMBER	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) <b>Naval Medical Research Center Silver Spring, Maryland, 20910</b>				8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) <b>Naval Medical Research Center 503 Robert Grant Avenue Silver Spring, MD 20910-7500</b>				10. SPONSOR/MONITOR'S ACRONYM(S)	
				11. SPONSOR/MONITOR'S REPORT NUMBER(S)	
12. DISTRIBUTION/AVAILABILITY STATEMENT <b>Approved for public release, distribution unlimited</b>					
13. SUPPLEMENTARY NOTES					
14. ABSTRACT					
15. SUBJECT TERMS					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT <b>SAR</b>	18. NUMBER OF PAGES <b>6</b>	19a. NAME OF RESPONSIBLE PERSON
a. REPORT <b>unclassified</b>	b. ABSTRACT <b>unclassified</b>	c. THIS PAGE <b>unclassified</b>			

## INTRODUCTION

Between May and October 2002, a number of febrile cases, including two deaths, were reported in an area around the town of Sapillica in northern Peru. A joint investigation with representatives of the Peruvian Ministry of Health and Naval Medical Research Center Detachment, Lima, documented a high prevalence of antibodies to *Rickettsia* and *Leptospira* species.<sup>1</sup> During the investigation nucleic acid preparations extracted from blood samples obtained from febrile humans and peri-domestic rodents, and from fleas and ticks from individual domestic and wild-trapped animals, were determined by PCR as specific for four rickettsial genes, including: the *Rickettsia* genus-specific 17-kDa outer membrane antigen gene (17-kDa gene), the citrate synthase gene (*gltA*), and the outer membrane protein A (*ompA*) and B (*ompB*) genes to contain evidence for the presence of SFG rickettsiae.<sup>2</sup> Initially, all samples were tested for presence of the 17-kDa gene, and if detected determination of whether the amplicon was from a typhus or spotted fever group rickettsia was made. Prevalence as determined from these tests demonstrated that 24% of febrile humans (4/17), 8.3% of trapped rodents (2/24), 1.6% of flea pools (1/59), and 25% of ticks (4/16) were positive for SFG agents. Further analysis of the molecular isolates utilizing PCR specific for the *gltA*, *ompA*, and *ompB* genes and sequencing the subsequent amplicons determined the phylogenetic relationships of the molecular isolates. In addition to detecting *R. felis* in flea triturates, we identified a new SFG family member molecularly isolated from two tick samples (tick 124 and tick 163).<sup>2</sup> Due to the more than 3% difference in sequence identity between a 2,484 bp sequence segment of *ompB* from the novel agent and those SFG rickettsiae sequences listed in GenBank, it was proposed that the new agent be named *Rickettsia andeanae*. Herein, we describe further the characterization of *Candidatus Rickettsia andeanae* based upon phylogenetic analysis of five rickettsial genes. Four of the genes have been previously assessed (17-kDa gene 434 bp; *gltA* 381 bp; *ompA* 540 bp; and *ompB* 2,484 bp)<sup>2</sup> however, for this project larger amplicons from 17-kDa gene (483 bp), *gltA* (1185 bp) *ompA* (1,589 bp), and *ompB* (4,839 bp), and 2,634 bp of *sca4* (120 kDa cytoplasmic protein gene; geneD) were analyzed and the results described in this report confirmed *Candidatus Rickettsia andeanae* inclusion within, and distance and uniqueness from other known SFG rickettsiae.

## MATERIALS AND METHODS

Two molecular isolates of *Candidatus Rickettsia andeanae* from two ticks (*Amblyomma maculatum* and *Ixodes boliviensis*) collected from domestic horses living in two separate locations in Northern Peru (Coletas and Naranjo)<sup>2</sup> were evaluated by multilocus sequence typing to determine their relationship to other rickettsiae. Following amplification by standard polymerase chain reaction (PCR) of the 17-kDa gene and nested PCR amplification of *ompB*, *ompA*, and *sca4*, nucleic acid-derived from the two isolates, amplicons were purified and sequenced on an automated ABI Prism 3100 gene analyzer (Applied Biosystems, Foster City, CA) similar to that described previously.<sup>2</sup> However, new primers were used for PCR and sequencing in this study (TABLE 1). The sequences (both forward and reverse) were assembled with

TABLE 1. Primers for PCR, nested PCR, and sequencing

Target Gene	Primer	Sequence (5'-3')	Amplicon Size (bp)
17-kDa antigen gene	Rp17kF1 <sup>ac</sup>	AATGAGTTTTATACTTTACAAAAT-TCTAAAAACCA	532
	Rr1175F <sup>ac</sup>	GCTCTTGCAACTTCTATGTT	434
<i>gltA</i>	Rr2608R <sup>ac</sup>	CATTGTTTCGTCAGGTTGGCG	1254
	CS1dF <sup>abc</sup>	ATGACTAATGGCAATAATAA	
	CS1273R <sup>a</sup>	CATAACCAGTGTAAGCTG	
	CS1234R <sup>bc</sup>	TCTAGGTCTGCTGATTTTTTGTTC	
	Rp CS877F <sup>bc</sup>	GGG GGC CTG CTC ACG GCG G	
<i>ompA</i>	Rp CS1258R <sup>bc</sup>	ATTGCAAAAAGTACAGTGAACA	382
	190-3588F <sup>a</sup>	AACAGTGAATGTAGGAGCAG	1630
	190-5238R <sup>abc</sup>	ACTATTAAAGGCTAGGCTATT	
	RompA1F <sup>bc</sup>	GAATAACATTACAVGCGYGGAGGAAG	
	RrA657R <sup>c</sup>	TATTTGCATCAATCSYATAAGWA	
<i>ompB</i>	RhoA4336F <sup>c</sup>	AGTTCAGGAAACGACCGTA	4887
	RompB11F <sup>a</sup>	ACCATAGTAGCMAGTTTTGCAG	
	RompB4887R <sup>ab</sup>	AGAGTACCTTGATGTGCRGTATAYT	
	RompB2409F <sup>bc</sup>	CCGTAACATTAACAAACAAGCTG	
	RompB2553R <sup>bc</sup>	GAATTTTCAAAGCAATYGTATCAGT	
	RAK1009F <sup>bc</sup>	ACATKGTTATACARAGTGYTAATGC	2479
	RAK1452R <sup>bc</sup>	SGTTAACTTKACCGYTTATAACTGT	2553
	120-2788F <sup>c</sup>	AAACAATAATCAAGGTACTGT	3069
	RompB3521F <sup>c</sup>	GATAATGCCAATGCAAATTTTCAG	
	RompB4224F <sup>c</sup>	ACCAAGATTATAAGAAAGGTGATAA	
	120-4346R <sup>c</sup>	CGAAGAAGTAACGCTGACTT	
	RompB3637R <sup>c</sup>	GAAACGATTACTTCCGGTTACA	
	RompB3008R <sup>c</sup>	AATATCGCTGACGGTCAAGGT	
	120-607F <sup>c</sup>	AATATCGCTGACGGTCAAGGT	
	120-807R <sup>c</sup>	CCTTTTAGATTACCGCCTAA	
	RR1595F <sup>c</sup>	GCCGGAGTTGTCCAATTATCA	
	RompB1902R <sup>c</sup>	CCGTCATTTCCAATAACTAACTC	
	RompB2012F <sup>c</sup>	GCCGGTACAAATTTAGGTAGTG	
<i>sca4</i>	D1f <sup>abc</sup>	ATGAGTAAAGACGGTAACCT	
	D3069f <sup>abc</sup>	TCAGCGTTGTGGAGGGGAAG	
	RrD749F <sup>abc</sup>	TGGTAGCATTAAGAGCTGATGG	
	RrD2685R <sup>abc</sup>	TTCAGTAGAAGATTTAGTACCAAAT	
	RrD928F <sup>bc</sup>	ATTTATACACTTGCGGTAACAC	
	RrD1826R <sup>bc</sup>	TCTAAATKCTGCTGMATCAAT	
	RrD1713F <sup>c</sup>	CTCTGAATTAAGCAATGCGGAAA	
	T5DR1 <sup>c</sup>	CTGATAAAGCTGTAGCTGCATTA	
	D2338f <sup>c</sup>	GATGCAGCGAGTGAGGCAGC	
	D928f <sup>c</sup>	AAGCTATTGCGTCATCTCCG	

<sup>a</sup>Primers used for PCR amplification; <sup>b</sup>Primers used for nested PCR amplification; <sup>c</sup>Primers used for sequencing.

TABLE 2. *Candidatus Rickettsia andeanae* sequence identity with closest neighbors

Gene	Tick Sample	Sequence Length (bp)	Length of ORF to <i>R. conorii</i>	Position to ORF	Results (percent identity with closest neighbors)	Fournier and colleagues Cutoff Values <sup>a</sup>
17-kDa	Tick124	422		43~464	98.49% to <i>R. rickettsii</i> ; 98.28% to <i>R. sibirica</i> and <i>R. conorii</i> ; 98.06% to <i>R. rhipicephali</i> ; 97.84% to <i>R. parkeri</i>	<sup>b</sup>
	Tick163	483	480	M30~464		
<i>gltA</i>	Tick124	354			99.42% to <i>Rickettsia</i> sp. DnS14; 99.34% to R.sp. RpA4, <i>R. sibirica</i> and R.sp. MC16, R.sp. BJ-90; 99.26% to <i>R. parkeri</i> , <i>R. slovac</i> a 99.17% to <i>R. conorii</i> , 99.12% to <i>R. aeschlimannii</i> ; 99.09% to <i>R. rhipicephali</i>	99.9%
	Tick163	1185	1308	810~1163 45~1229		
<i>ompA</i>	Tick124 and Tick163	1598	6066	2880~4464	98.03% to R.sp. RpA4; 97.90% to R.sp. DnS28; 97.84% to R. sp. DnS14; 97.11% to <i>R. rhipicephali</i> and <i>R. massiliae</i>	98.8%
<i>ompB</i>	Tick124 and Tick163	4839	4968	89~4927	97.22% to <i>R. aeschlimannii</i> ; 97.20% to <i>R. rhipicephali</i> ; 97.10% to R. sp. Bar 29; 97.08% to <i>R. massiliae</i> ; 96.10% to <i>R. slovac</i> a	99.2%
<i>sca4</i>	Tick124 and Tick163	2634	3081	39~2672	97.8% to <i>R. massiliae</i> ; 97.7% to <i>R. rhipicephali</i> and <i>R. slovac</i> a; 97.6% to <i>R. aeschlimannii</i> ; 97.5% to <i>R. conorii</i>	99.3%

Sequencher 4.0 software (Gene Codes Corp, Ann Arbor, MI) and compared to those downloaded from GenBank. Phylogenetic analyses (Neighbor Joining Best Tree and Bootstrap methods) were performed with MacVector 7.0 software (Accelrys, Inc, San Diego, CA).

RESULTS

17-kDa Gene and *gltA*

Comparing sequences of the *Candidatus Rickettsia andeanae* 17-kDa gene and *gltA* fragments with GenBank sequences indicate that *Candidatus Rickettsia andeanae* belongs within the genus *Rickettsia* by showing that the 483 bp 17-kDa gene segment had >97.5% identity with five SFG rickettsiae, but with no more then 98.5% identity with any rickettsial agent (TABLE 2). Similarly, the 1,185 bp *gltA* sequence fragment of *Candidatus Rickettsia andeanne* had 99.00–99.42% identity to ten SFG

rickettsiae, but no identity with any rickettsial agent >99.42% (TABLE 2). Thus, the sequencing data from the 17-kDa gene and *gltA* segments indicate that *Candidatus Rickettsia andeanae* is a rickettsial agent but not identical with any agent currently listed within GenBank.

#### ompA

Due to presence of *ompA* among SFG rickettsiae, and the lack of a complete *ompA* among TG rickettsiae, *R. bellii* and *R. canadensis*, the production of a PCR amplicon utilizing *ompA* specific primers corroborates the *gltA* data that suggests that *Candidatus Rickettsia andeanae* is closely related to SFG rickettsiae. In addition, phylogenetic analysis of a 1598 bp segment of *ompA* indicates that *Candidatus Rickettsia andeanae* has 98.03% identity with its closest neighbor, *Rickettsia* sp. RpA4, within the SFG (TABLE 2).

#### ompB

Initial data from the phylogenetic analysis of *ompB* sequence (2,484 bp) from *Candidatus Rickettsia andeanae* showed that it was  $\geq 3\%$  divergent from other SFG rickettsiae.<sup>2</sup> In this investigation an *ompB* fragment almost twice as large (4,839 bp) was sequenced and evaluated for its relationship to other rickettsial (and non-rickettsial) sequences. It was determined that the larger segment also showed an approximately 3% divergence from the *ompB* sequences of the closest relations, which were determined to be SFG rickettsiae (TABLE 2). The most closely related rickettsiae based upon comparison of the 4,839 bp fragment of *ompB* were *R. aeschlimannii* (97.22% identity), *R. rhipicephali* (97.20% identity), and *R. sp.* Bar 29 (97.10% identity). The *Candidatus Rickettsia andeanae ompB* sequence showed only 87.7 and 87.3% identity with *R. prowazekii* and *R. typhi*, respectively.

#### sca4

Sequence comparisons of part or all of *sca4* have recently been described as a tool for rickettsial phylogenetic analysis because the presence of the gene among the various species of *Rickettsia* and due to its moderate sequence variability among the *Rickettsia*.<sup>3</sup> Blast search with the 2,634 bp *sca4* segment sequence from *Candidatus Rickettsia andeanae*, showed that it was most closely related to SFG rickettsiae *R. massiliae* (97.8%), *R. rhipicephali* and *R. slovaca* (97.7%). Similar to the sequence comparison data discerned with *ompA* and *ompB*, *Candidatus Rickettsia andeanae sca4* sequence shows a close relationship to SFG rickettsiae, but is divergent enough to be considered a unique rickettsial species (TABLE 2).

## DISCUSSION

Multilocus sequence typing (MLST) for the characterization of pathogenic and non-pathogenic microorganisms utilizes the sequences from several gene segments to characterize the genetic makeup of an isolate.<sup>4</sup> The sequences are compared to those already obtained for similar and dissimilar organisms. From the comparison a determination is made as to the relationship between the new isolate and known iso-

lates. This is very helpful in ascertaining the identity of an isolate(s) during an outbreak investigation, historical classification, or taxonomic relationships and has been used to characterize rickettsial agents.<sup>3,5</sup> In this investigation PCR amplified segments from five well known genes were used to validate the divergence of the novel isolate, *Candidatus Rickettsia andeanae*, from those rickettsiae previously characterized utilizing the criteria proposed by Fournier and colleagues.<sup>3</sup> The results presented herein confirm our earlier report<sup>2</sup> by extending the sequences assessed for four genes (17 kDa gene, *gltA*, *ompA*, *ompB*) and newly assessing the sequence of *sca4*. Collectively these results show the uniqueness of *Candidatus Rickettsia andeanae*, a novel agent that has been molecularly isolated from two different ticks collected from two different horses from two locations in northern Peru. However, an isolate has not been cultured in the laboratory and therefore the pathogenicity of this agent has not been ascertained. Collaborators in Peru are actively pursuing a course for cultivating this agent in their laboratory.

### ACKNOWLEDGMENTS

The work reported herein was supported by the DoD GEIS program and its work unit number was 0000188M.0931.001.A0074. The opinions and assertions contained herein are the private ones of the authors and are not to be construed as official or reflecting the views of the Department of the Navy, the Department of Defense or the U.S. Government.

### REFERENCES

1. BLAIR, P.J., G.B. SCHOELER, C. MORON, *et al.* 2004. Evidence of rickettsial and leptospira infections in Andean Northern Peru. *Am. J. Trop. Med. Hyg.* **70**: 357–363.
2. BLAIR, P.J., J. JIANG, G.B. SCHOELER, *et al.* 2004. Characterization of spotted fever group rickettsiae in flea and tick specimens from northern Peru. *J. Clin. Microbiol.* **42**: 4961–4967.
3. FOURNIER, P-E., J.S. DUMLER, G. GILBERT, *et al.* 2003. Gene sequence-based criteria for identification of new *Rickettsia* isolates and description of *Rickettsia heilongjiangensis* sp. nov. *J. Clin. Microbiol.* **41**: 5456–5465.
4. MAIDEN, M.C.J., J.A. BYGRAVES, E. FEIL, *et al.* 1998. Multilocus sequence typing: a portable approach to the identification of clones within populations of pathogenic microorganisms. *Proc. Natl. Acad. Sci. USA* **95**: 3140–3145.
5. ISHIKURA, M., S. ANDO, Y. SHINAGAWA, *et al.* 2003. Phylogenetic analysis of spotted fever group rickettsiae based on *gltA*, 17-kDa, and *rOmpA* genes amplified by nested PCR from ticks in Japan. *Microbiol. Immunol.* **47**: 823–832.